A FIM-based Long-Term in-vial Monitoring System for Drosophila Larvae

A FIM-based Long-Term in-vial Monitoring System for Drosophila Larvae

Dimitri Berh‡, Benjamin Risse‡, *Member, IEEE*, Tim Michels, Nils Otto, Xiaoyi Jiang*, *Senior Member, IEEE*, Christian Klämbt

***Abstract*—Drosophila larvae are an insightful model and the automated analysis of their behaviour is an integral readout in behavioural biology. Current tracking systems, however, entail a disturbance of the animals, are labour-intensive, and cannot be easily used for long-term monitoring purposes. Here we present a novel monitoring system for Drosophila larvae which allows us to analyse the animals in cylindrical culture vials. By utilizing frustrated total internal reflection in combination with a multi-camera/microcomputer setup we image the complete housing vial surface and thus the larvae for days. We introduce a calibration scheme to stitch the images from the multi-camera system and unfold arbitrary cylindrical surfaces to support different vials. As a result, imaging and analysis of a whole population can be done implicitly. For the first time, this allows to extract long-term activity quantities of larvae without disturbing the animals. We demonstrate the capabilities of this new setup by automatically quantifying the activity of multiple larvae moving in a vial. The accuracy of the system and the spatio-temporal resolution is sufficient to obtain motion trajectories and higher level features like body bending. This new setup can be used for in-vial activity monitoring and behavioural analysis and is capable of gathering millions of data points without both disturbing the animals and additional labour time. In total we have analysed $107,671$ frames resulting in $8,650$ trajectories, which are longer than 30 seconds, and obtained more than 4.2×10^6 measurements.**

***Index Terms*—Drosophila Larvae, FIM, FTIR, Imaging Technique, Locomotion Analysis, Long-Term, Microcomputer, Multi-camera, Unfolding, Stitching**

I. Introduction

THE importance of behavioural studies of model organisms has significantly increased in recent biological research. Complex experimental constructions and high-throughput experiments are only achievable by computer-aided image acquisition and analysis techniques [1]–[11].

To obtain meaningful interpretation and understanding of the behavioural phenotype of an animal, experiments usually have to be repeated multiple times under the same environmental conditions. Most importantly, the animals have to be taken out of their rearing habitat, cleaned (i.e. by removing dirt and other artefacts), and are afterwards placed in a new environment. These steps may disturb the animals and thereby lead to falsified behaviour output. Moreover, these steps are time-consuming and repetitive for the experimenter.

In current research concerning social interactions, day rhythm, or age related deficits, long-term automatic monitoring and behaviour analysis of freely moving animals is of main interest. This concerns mice [12]–[14], *Caenorhabditis elegans* [15]–[17], as well as Drosophila flies [18]–[22]. However, to the best of our knowledge there is no system available for monitoring the behaviour of Drosophila at the larval states in culture vials for a long time period.

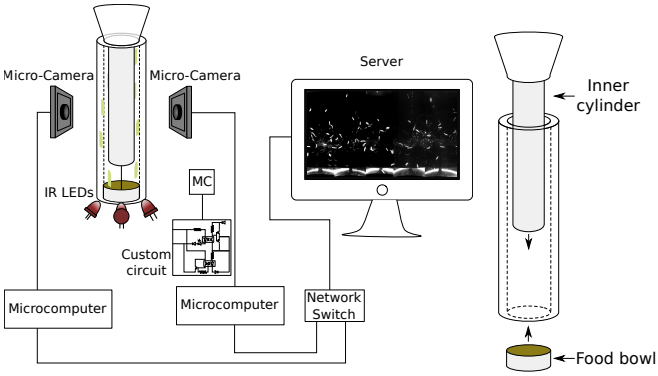


Fig. 1. Overview of the proposed setup. (a) We use microcomputers and micro-cameras to monitor *Drosophila melanogaster* larvae in cylindrical rearing vials for days by utilizing the frustrated total internal reflection effect. The microcomputers are synchronized, act as autonomous devices, and are connected in a client-server-architecture to a central server. Preprocessed images from the microcomputers are stitched together and further processed on the server. (b) Custom-made vial with the main components such as the inner cylinder and the food bowl. For details see text.

A. Related Work

To analyse a large number of genotypes in a high-throughput manner animals have to be monitored for a long time period. Jhuang *et al.* [12] have suggested an approach for monitoring freely moving mice. Here an algorithm was trained with manually annotated image examples of behaviours of interest of home-caged mice. This trained system was then used to automatically annotate behaviours of interest in new recordings containing hours of material for one freely moving mouse in a cage. Ohayon *et al.* [13] have proposed a method for automated multi-day tracking of freely moving mice. Each

This work has been partially supported by the Deutsche Forschungsgemeinschaft, DFG EXC 1003 Cells in Motion - Cluster of Excellence, Münster, Germany (PP-2014-05, FF-2016-06, CIM-IMPRS fellowship to N.O.).

D. Berh and X. Jiang are with the Department of Mathematics and Computer Science, University of Münster, Münster, Germany, e-mail: d.berh@uni-muenster.de, xjiang@uni-muenster.de

B. Risse is with the School of Informatics, University of Edinburgh, Edinburgh, Scotland, UK, e-mail: brisse@inf.ed.ac.uk

T. Michels is with the Department of Computer Science, University of Kiel, Kiel, Germany, e-mail: tmi@informatik.uni-kiel.de

N. Otto and C. Klämbt are with the Department of Behaviour and Neuronal Biology, University of Münster, Münster, Germany, e-mail: nils.otto@uni-muenster.de, klaembt@uni-muenster.de

Authors marked by ‡ contribute equally and the corresponding author is marked by *

mouse was marked by a unique fur pattern using hair bleach and an image classifier was trained based on samples of the unique patterns of each mouse. Weissbrod *et al.* [14] have presented another automated long-term tracking and social behaviour analysis system. By combining video and radio frequency techniques (RFID) the system is able to obtain behavioural profiles of both individuals and groups. In practice, mice with two implanted RFID microchips were placed in a testing enclosure which was equipped with a custom-made array of circular RFID receivers horizontally placed beneath the arena floor. Additionally the position of the animals was recorded by a video camera.

Simonetta *et al.* [15] have proposed a new method for long-term recordings of *Caenorhabditis elegans* based on infrared microbeam scattering. The measurement technique utilizes an infrared light source which is mounted beneath the tracking stage. The light crosses a plate with an array of 100μm wide micro holes before it can pass through the worm culture receptacle and finally gets measured by a phototransistor. Yu *et al.* [16] have described a new automated monitoring technique for multiple worms of any development stage for several days. In practice, an array of plano-concave glass wells containing one worm (or a pair for mating experiments) was illuminated using dark field illumination and recording was done by a camera mounted on top of the array.

Besides the aforementioned model organisms, Drosophila is widely used. Several automated activity monitoring and locomotion analysis studies are related to Drosophila flies. Among all setups, the commercially available *TriKinetics* [23] system is one of the most popular. . Here a single fly is placed in a closed glass tube with food and air supply. Each tube is crossed by up to 16 infrared light beams and a detector records how often the fly interrupts one of the beams in a user defined time period by walking through it. Gilestro [21] has suggested an improvement of the aforementioned setup for analysing sleep behaviour in *Drosophila melanogaster*. Instead of using separate infrared emitter-detector pairs, a commercially available video camera was used to video-record multiple arenas of different shapes enclosing one fly for days. Dankert *et al.* [18] have introduced a method for automatically measuring aggression and courtship. They have used a setup with up to 4 *double arenas*, where one pair of flies was enclosed per arena. Inan *et al.* [20] have introduced a setup to image Drosophila flies in a petri dish which was illuminated from below by a uniform light source. A miniature, low-cost CMOS video camera was used to obtain a video stream which was further processed to extract the average light changes.

All of the previously mentioned methods and setups have several drawbacks when it comes to monitoring Drosophila larvae in habitats such as cylindrical culture vials, since it is not possible to use bleaching or RFID. Furthermore, monitoring only one or two larvae in one enclosure is not desirable if community behaviour is of interest.

Recently we have introduced an imaging technique for small animals, called FIM (FTIR-based Imaging Method) [24]. This technique utilizes frustrated total internal reflection (FTIR) to measure the contact surface of the animals on the substrate. The acquired images have an unsurpassed foreground/background contrast. In subsequent studies we have shown that FIM in combination with the associated software FIMTrack is suitable for a wide range of biological applications as well as a wide range of organisms [24]–[29].

Here, we propose an evolution of our previous work by extending our FIM setup into an automatic FIM-based long-term monitoring system to minimize the aforementioned influences and to speed up the experiments. An overview of our setup is sketched in Fig. 1. Additionally Fig. 2b and Fig. 2c give an impression of the current design. The utilization of the FTIR effect in combination with micro-cameras and microcomputers allows the experimenter to monitor third instar larvae in culture vials over several hours up to days without disturbing them. Furthermore, the time-consuming task of preparing each individual run of the experiment is reduced to a minimum. Our system is easy to use and since no specialized hardware is employed, the system is inexpensive and scalable.

B. Organization

This paper is structured as follows. Section II gives a brief overview of the components of the setup. In Section II-A the hardware design in combination with the physical principle of the FTIR effect is described whereas Section II-B provides a more detailed overview of the software components. The necessary algorithmic steps are described in Section III. Particularly the unfolding process of the images and the stitching step are explained in Section III-B and Section III-C in detail. The quality of the unfolding algorithm (Section IV-A), the performance (Section IV-B), the contrast and the overall image quality (Section IV-C), and first monitoring results (Section IV-E) are examined in the next section. A discussion and a final conclusion are given in Section V and Section VI.

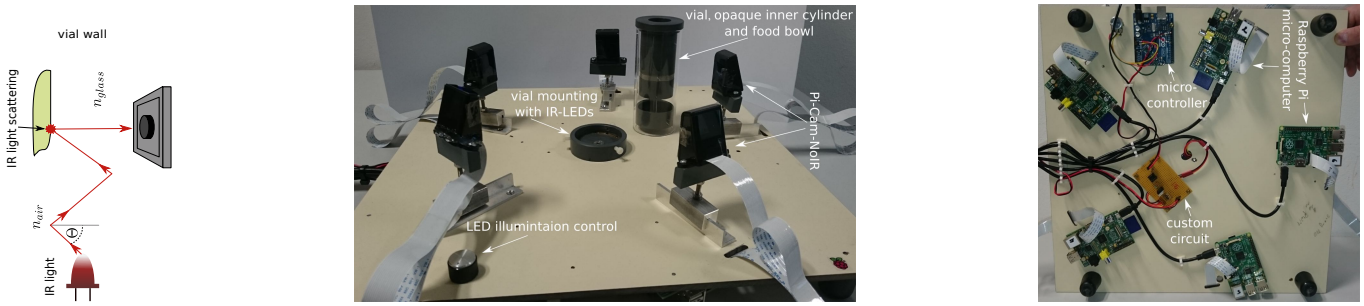
II. System Overview

A. Hardware

Usually animals like Drosophila larvae are reared in so-called *culture vials* which normally have a cylindrical shape. The material of such vials varies from plastic over acrylic glass to glass. We use custom vials made out of borosilicate glass. In [24], [27] it was shown that those materials are suitable for the frustrated total internal reflection (FTIR) effect. Here we will provide a short summary of the physical principles facilitating this effect (Fig. 2a).

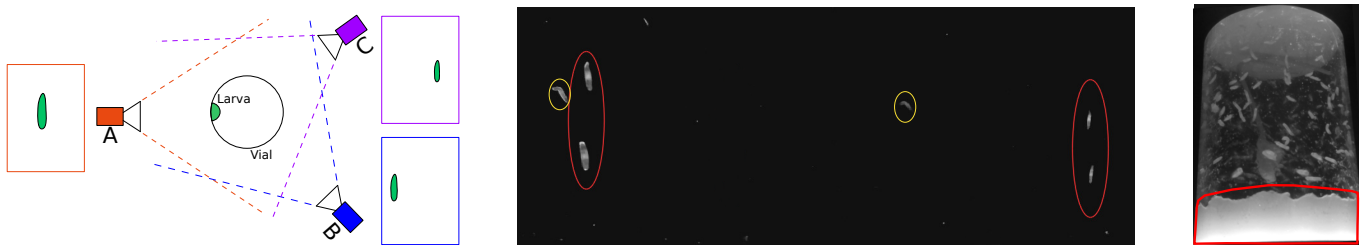
Given the approximated refractive indices of air $n_{air} \approx 1.00$ and borosilicate glass $n_{glass} \approx 1.52$ [30] the critical angle at the glass/air boundary according to Snell's law is given by $\Theta = \arcsin\left(\frac{n_{air}}{n_{glass}}\right) = \arcsin\left(\frac{1.00}{1.52}\right) \approx 41.14°$. All infrared (IR) light rays with angles of incidence above Θ are completely reflected at the glass/air boundary since $n_{glass} > n_{air}$, which facilitates the total internal reflection effect. In our setup the light rays are directed into the borosilicate glass at the cylinder's bottom edge. When a larva is touching the vial wall, the IR light can enter the semi-translucent animal body and gets scattered by larva's tissue. Among others, the scattering process produces light rays with angles of incidence below the critical angle Θ. These rays of light are not completely

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBME.2016.2628203, IEEE Transactions on Biomedical Engineering



(a) (b) (c)

Fig. 2. Proposed setup. (a) Sketch of the frustrated total internal reflection effect in a culture vial. (b) Top view of the current prototype with 5 cameras, vial mounting and the vial. The food bowl and the inner cylinder are indicated. IR-LEDs are placed in the vial mounting in order to direct the light into the glass vial from underneath. (c) All electronic components are placed underneath. For details see text.



(a) (b) (c)

Fig. 3. Challenges in our hardware setup: (a) Sketch of the *ghost* problem. The larva facing camera **A** is clearly visible in the associated image. Since IR light can pass through the larval body, the larva is also visible in cameras **B** and **C**. This signal is usually less bright than the scattered light. (b) Real life example of the ghosts. The ghosts of the larvae marked by red and yellow circles on the left side are marked with the same colours on the right side. (c) *Glowing food*: Since the food (red area) is touching the vial wall it induces the FTIR effect. Thus, IR light gets scattered and illuminates the vial from inside.

reflected any more. They are *frustrated* and can pass through the glass leading to the FTIR effect. This light is captured by surrounding cameras. Therefore the captured images show a hight contrast since only light scattered by larvae is visible.

To induce the FTIR effect 12 IR LEDs (Osram Components SFH 4350) with a dominant wavelength of 860nm are mounted at the bottom of the vial and connected to a custom-made circuit board. To precisely adjust the intensity of the LEDs additionally an Arduino Uno microcontroller (MC) is connected to the circuit in combination with a control knob allowing to adjust the intensity of the IR LEDs without plugging the MC into a computer. The experimenter can easily turn the LEDs on and off as well as control their intensity between 0% and 100%. The MC translates the chosen intensity by setting the pulse-width modulation (PWM) to a value in the range of $[0, 255]$, where 0 means no power and 255 indicates the maximum power the LEDs are able to provide.

To monitor the whole cylindrical surface of a vial a multi-camera setup is used. For a modular and flexible setup, each of the cameras (Raspberry Pi HD NoIR v1.3 with 5MP) is connected to a standalone microcomputer (Raspberry Pi B+ with a Broadcom BCM2835 700MHz ARM11 CPU, 512 MB SDRAM, and a 10/100 Mbit/s Ethernet connection) and is mounted on a guide rail approximately 17cm away from the centre of the vial mounting. Furthermore, the cameras and microcomputers act as autonomous devices. Based on this design, it is possible to replace and configure each camera and microcomputer separately. The cameras and microcomputers are connected to a central server (Dell Optiplex 9020 with an Intel Core i7-4790 3.60GHz CPU, 16 GB DDR3-1600 RAM) in a client-server-architecture using an 8-Port Gigabit Switch (TEG-S80g).

Since the animals are semi translucent the IR light can partially pass through the animals' body and since a multi-camera setup is used where the cameras are placed around the vial, the IR light transmitted through the larval body is captured by cameras which are imaging the dorsal part of the larvae. This signal leads to so-called *ghosts*: A single larva is visible multiple times resulting in an increased number of recorded animals in the tracking area. This is sketched in Fig. 3a and a real life example is given in Fig. 3b. To remove the ghosts we physically avoid the signal from all larvae not facing the camera ventrally by inserting an inner cylinder which is made of an opaque material as a light barrier (Fig. 2b, results can be found in Section IV-D). Furthermore, food for the larvae is usually directly contained within the culture vials. However, this causes another challenge: if the food is touching the vial wall, the FTIR effect occurs at the glass/food boundary. Thus, the IR light gets scattered and illuminates the inner part of the cylinder leading to a decrease of the foreground/background contrast (Fig. 3c). To prohibit this effect we have designed a food bowl made of an opaque materiel (Fig. 2b) which is mounted underneath the inner cylinder in order to uncouple the food and the IR light.

In theory, all vials with adequate optical properties can be used since our unfolding approach is not relying on any

specific vial diameter (see Section III). However, the diameter of the food bowl must be adjusted accordingly to be inserted into the cylinder. To optimize the fly handling protocol we utilized a custom-made vial which is open on both sides. This allows for an easy removal of the flies after oviposition so that the food bowl contains the eggs (Fig. 1b). The inner cylinder is inserted from the top and the vial can be mounted as shown in Fig. 2b.

B. Runtime Environment

Both the client and the server are implemented in C++ utilizing the OpenCV library (v2.4). To be able to access the camera connected to the Raspberry Pi directly on the hardware level the Multi-Media Abstraction Layer API provided by Broadcom Europe Ltd. and redistributed by the Raspberry Pi (Trading) Ltd. is used [31]. Moreover, the server employs the Qt (v5.5) framework for the graphical user interface. For communication purposes between the clients and the server we use a custom-tailored network protocol which allows the experimenter to adjust all relevant camera parameters as well as to start or stop the image acquisition and transmission process directly from the server as a central node. Additionally, camera synchronization is implemented on the software level using the aforementioned network communication protocol.

Through utilization of the FTIR effect the produced images have bright objects as foreground and an almost black background. Therefore a grey level image will be sufficient to encode all relevant informations. To reduce computation time on the clients and to speed up the transmission process from the clients to the central server we capture the images in the *YUV* format and discard the U and V components to get a one-channel grey level image, since the Y component is sufficient for describing the brightness of the image.

III. Methods

The curved geometry of the cylindrical culture vial causes perspective distortions in the images. In order to provide an accurate analysis of small animals like Drosophila larvae, an exact unfolding of the vial surface is crucial prior to further analysis. Additionally, due to the use of a multi-camera setup, a stitching of the individual images from each camera is desirable to provide an overall view of the surface. The process results in three main steps, namely *keypoint detection and sorting*, *unfolding of a cylindrical surface* and *image stitching*. It should be noted, that these steps have to be done only once. After performing these steps the experimenter is able to remove the calibration pattern and to insert the food bowl together with the inner cylinder without moving the vial.

A. Keypoint Detection and Sorting

To precisely image and analyse animals in a cylindrical vial the first step is to establish correspondences between the three dimensional geometry of the cylinder and the camera images. Therefore, keypoint detection is performed by using a checkerboard pattern (Fig. 4) in combination with a corner detector. Since the vial wall has a certain thickness we place the checkerboard pattern on the inner side of the vial in order to consider the same surface the larvae will crawl on and thus to achieve a higher accuracy.

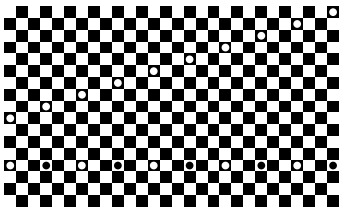


Fig. 4. Checkerboard pattern for keypoint detection and stitching. The circles on the pattern are important to provide a unique pattern for the image stitching step in Section III-C.

Currently the checkerboard corner detector implemented in OpenCV is used. Additionally, the user has the possibility to remove false or to add missing keypoint detections.

To be able to establish correspondences between the keypoints in the images and the markers on the cylinder it is necessary to ensure a specific order of these keypoints. Given a set of keypoints $\mathcal{K}$ on the image plane (Fig. 5a), we first identify the four outer corners $k_1, \ldots, k_4$ by utilizing the fact that for every other point $k \in \mathcal{K}$ it is possible to find neighbour points $\ell_1, \ell_2 \in \mathcal{K}$ where $\angle \ell_1 k \ell_2 \approx 180°$. Thus, for every detected keypoint k we search for points ℓ_1, ℓ_2 such that

$$|\angle \ell_1 k \ell_2 - 180°| < t_\alpha \tag{1}$$

where ℓ_1, ℓ_2, k have to be pairwise different. The value t_α is a threshold defining the maximum turning angle between the line segments $\overline{\ell_1 k}$ and $\overline{k \ell_2}$ so that the points ℓ_1, k, ℓ_2 are interpreted as approximately lying on a straight line. A keypoint k is one of the four outer corner points if no points ℓ_1, ℓ_2 satisfying (1) can be found.

At the end of this step we arrange the four outer corners $k_1, \ldots, k_4$ in a clockwise order around their geometric midpoint given by $\frac{k_1+k_2+k_3+k_4}{4}$. The final order of the four outer corners is illustrated in Fig. 5a.

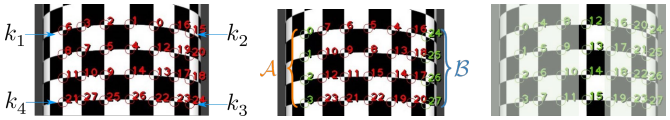
Next, we compute two sets of keypoints $\mathcal{A}, \mathcal{B} \subset \mathcal{K}$. The set $\mathcal{A}$ contains all keypoints lying approximately on the line connecting k_1 and k_4 including k_1 and k_4 and $\mathcal{B}$ contains all keypoints lying approximately on the line connecting k_2 and k_3 including k_2 and k_3, respectively. Keypoints in $\mathcal{A}$ and $\mathcal{B}$ are sorted according to the y-coordinate. Fig. 5b illustrates the first sorting step.

Afterwards all previous mentioned steps including finding the outer four corners $k_1, \ldots, k_4$ are performed on the set $\mathcal{K} \setminus \{\mathcal{A} \cup \mathcal{B}\}$ recursively as long as there are at least two keypoint columns left. Otherwise the remaining keypoints are sorted according to the y-coordinate without searching for the four outer corners $k_1, \ldots, k_4$. Once the recursion ends all sorted keypoint subsets are merged.

B. Unfolding of Cylindrical Surface

Several algorithms deal with the subject of unfolding a cylindrical surface [32], [33]. Since our setup is compact, static, and it is possible to use a calibration pattern easily we introduce a more suitable and simplified method.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBME.2016.2628203, IEEE Transactions on Biomedical Engineering



(a) (b) (c)

Fig. 5. Illustration of the keypoint sorting algorithm. The unfolding process is constrained to a more central area since the distortion gets worse at the visible edges of the vial. (a) Unsorted detected keypoints. The four outer corner points $k_1, \dots k_4$ are marked. (b) Keypoint indices after performing the first sorting step. The sets $\mathcal{A}, \mathcal{B}$ are marked by keypoint identifier written in green. (c) Last step of the sorting algorithm. The identifiers of sorted keypoints are written in green and keypoints already sorted in previous steps are faded out.

A right circular cylinder surface $\mathcal{M} \subset \mathbb{R}^3$ with radius r and height h located at the origin can be parametrized by

$$\mathcal{M} = \{(r \cdot \cos(\phi), r \cdot \sin(\phi), z) \mid \phi \in [0, 2\pi], z \in [0, h]\} \quad (2)$$

Using a rotation matrix $A \in SO(3)$ and a translation vector $b \in \mathbb{R}^3$ a point on any arbitrary right circular cylinder surface can be expressed by

$$Ap + b, \;\; p \in \mathcal{M} \quad (3)$$

If the parameters A, b and the focal length of the camera f are known, the projection $q = (x, y, 0)$ of the point $Ap + b$ can be obtained by solving the equation

$$q = c + t \cdot (Ap + b - c) \quad (4)$$

where $c = (0, 0, -f)$ is the camera centre (Fig. 6). A priori the parameters A, b and f are unknown. However, (4) can be rewritten as

$$\begin{aligned}
x &= 0 + t(a_{11} r\cos(\phi) + a_{12} r\sin(\phi) + a_{13} z + b_1 - 0) \\
y &= 0 + t(a_{21} r\cos(\phi) + a_{22} r\sin(\phi) + a_{23} z + b_2 - 0) \\
0 &= -f + t(a_{31} r\cos(\phi) + a_{32} r\sin(\phi) + a_{33} z + b_3 + f) \\
\Rightarrow 0 &= f a_{11} r\cos(\phi) + f a_{12} r\sin(\phi) + f a_{13} z + f b_1 \\
&\quad - a_{31} r\cos(\phi) x - a_{32} r\sin(\phi) x - a_{33} z x - b_3 x - f x \\
\wedge\ 0 &= f a_{21} r\cos(\phi) + f a_{22} r\sin(\phi) + f a_{23} z + f b_2 \\
&\quad - a_{31} r\cos(\phi) y - a_{32} r\sin(\phi) y - a_{33} z y - b_3 y - f y \\
\Rightarrow &\underbrace{\begin{pmatrix} \alpha & \beta & z & 0 & 0 & 0 & -\alpha x & -\beta x & -z x & 1 & 0 & -x & -x \\ 0 & 0 & 0 & \alpha & \beta & z & -\alpha y & -\beta y & -z y & 0 & 1 & -y & -y \end{pmatrix}}_{=: D_{x,y,r,\phi,z}} \cdot v = 0
\end{aligned}$$

where $v = (f a_{11}, f a_{12}, f a_{13}, f a_{21}, f a_{22}, f a_{23}, a_{31}, a_{32}, a_{33}, f b_1, f b_2, b_3, f)^t$ and $\alpha := r\cos(\phi)$, $\beta := r\sin(\phi)$.

Since the geometry of the used cylinder and the checkerboard is known, the angle between two horizontally adjacent corners is given by

$$2\pi \cdot \frac{s_c}{U} = \frac{s_c}{r}$$

where $U = 2\pi r$ is the perimeter of the cylinder and s_c is the real word square size of the used checkerboard pattern (Fig. 6). Based on this knowledge it is possible to compute the associated point in cylindrical coordinates for every detected and sorted keypoint as depicted in Fig. 7. Without loss of generality the two dimensional surface coordinate system of the cylinder image can be placed at the corner $k \in \mathcal{K}$ which is the lowest and leftmost detected corner. Thus, for every point

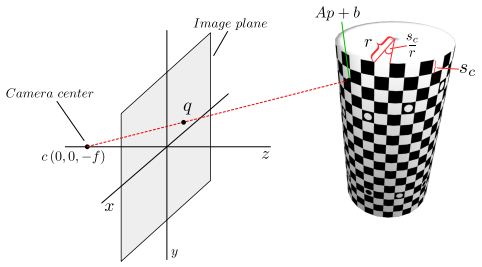
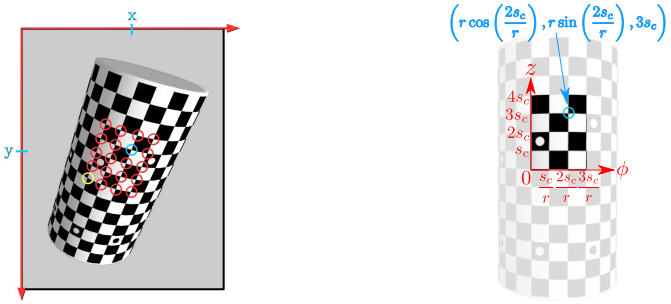


Fig. 6. Projection of a point $Ap + b$ on a cylinder onto the image plane in a standard pinhole camera model. Additionally, the properties of the coordinate system defined by the square size s_c of the checkerboard and radius r of the cylinder are depicted.



(a) (b)

Fig. 7. Corresponding point computation based on a detected 4×5 checkerboard pattern. (a) The lowest left detected corner $k \in \mathcal{K}$ is marked by yellow. (b) For the point k_i marked by blue we get $n_{i,c} = 2$ and $n_{i,r} = 3$.

$k_i \in \mathcal{K}$ the corresponding point p_i in cylindrical coordinates is given by

$$p_i = \left(r\cos\left(n_{i,c} \cdot \frac{s_c}{r}\right), r\sin\left(n_{i,c} \cdot \frac{s_c}{r}\right), n_{i,r} \cdot s_c\right)$$

where $p_i \in \mathcal{M}$, $n_{i,c}$ is the number of detected keypoint columns between k and k_i, and $n_{i,r}$ is the number of detected keypoint rows between k and k_i, respectively.

Given at least 7 point correspondences such that point $Ap_i + b$, $p_i \in \mathcal{M}$ on the cylindrical surface corresponds to point k_i on the image plane the following set of equations can be obtained

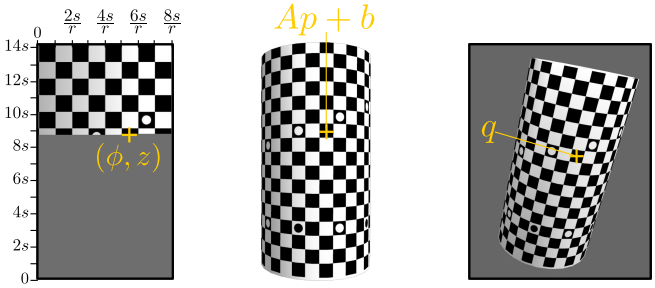
$$\begin{pmatrix} D_{x_1,y_1,r,\phi_1,z_1} \\ D_{x_2,y_2,r,\phi_2,z_2} \\ \vdots \end{pmatrix} \cdot v =: Dv = 0$$

To get a least squares solution Singular Value Decomposition [34] on the normal equation $D^t D v = 0$ can be used.

It should be noted that the least squares solution is obtained without the constraint $A^T A = I$ and thus the resulting matrix A is not necessarily in $SO(3)$. However, the accuracy of the algorithm is still sufficient (see Section IV-A).

In the next step we generate an arbitrarily sized rectangular (ϕ, h)-image $\mathcal{I}$ as depicted in Fig. 8 which is defined by both the square size s in pixels and the size of the detected checkerboard pattern. For every point in $\mathcal{I}$ we are able to compute the associated point on the cylindrical surface based on (2) and (3). By applying (4) we can obtain the projection point q on the image received from the camera. By bilinear interpolating the position of q, the grey value of the corresponding point on $\mathcal{I}$ is obtained. Fig. 8 illustrates the whole procedure.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBME.2016.2628203, IEEE Transactions on Biomedical Engineering



(a) (b) (c)

Fig. 8. Unfolding of a cylindrical surface. (a) Point coordinates (ϕ, z) on the (ϕ, h)-image $\mathcal{I}$ during construction of the unfolded image. (b) The three dimensional coordinate of the associated point on the cylindrical surface can be computed utilizing (2) and (3) in combination with the known parameters A and b. (c) Obtained image point q after projection of $Ap+b$ based on (4) and interpolation.

Since the cameras and the cylinder are static and to decrease the computation time of the unfolding step for each frame on the clients a *Lookup Table* is used. This table has the same size as $\mathcal{I}$ and contains for each point $(\phi, z) \in \mathcal{I}$ the associated projected point q.

C. Image Stitching

In contrast to image stitching algorithms for generating a panoramic view, in our setup we compute a horizontal and/or vertical offset and place the individual images from each camera side-by-side in a specific order. For computing the offsets we use the pattern given in Fig. 4. The different distances of the circles lying on the same vertical line allow an exact identification of the location of a single view with respect to the overall pattern.

In a given unfolded subview we first search for circles using Hough Transformation. By looking at the leftmost two vertically aligned circle midpoints $c^1 = (x, y_1)$ and $c^2 = (x, y_2)$ where $y_1 < y_2$ and by knowing the square size s the relative positions of these midpoints in the subpattern in square coordinates are given by $\tilde{c}^1 = \left(\lfloor \frac{x}{s} \rfloor, \lfloor \frac{y_1}{s} \rfloor\right)$ and $\tilde{c}^2 = \left(\lfloor \frac{x}{s} \rfloor, \lfloor \frac{y_2}{s} \rfloor\right)$, as depicted in Fig. 9. Additionally the vertical distance in squares between these midpoints is given by $d\left(\tilde{c}^1, \tilde{c}^2\right) = \tilde{c}^2_y - \tilde{c}^1_y$. Utilizing the fact that circle midpoints with a distance of 4 are located at the first (i.e. the 0th) column and that each additional circle midpoint pair has an offset of 3 squares to its predecessor in x-direction, the absolute column of the view in relation to the whole checkerboard is given by $c_{abs} = \left(d\left(\tilde{c}^1, \tilde{c}^2\right) - 4\right) \cdot 3$.

Next, we establish the absolute position of the subview by computing the upper left and the bottom right square index according to the global pattern. Since the base row with the alternating black and white circle sequence of the used checkerboard pattern is 13, the upper left square position can be estimated by $p^{ul} = \left(c_{abs} - \tilde{c}^2_x, 13 - \tilde{c}^2_y\right)$. Furthermore, given the dimensions of the detected checkerboard in the subview by n_x and n_y, we compute the bottom right square in absolute coordinates by $p^{br} = \left(p^{ul}_x + n_x, p^{ul}_y + n_y\right)$. Since the positions of the cameras as well as of the cylinder are fixed this procedure has to be done only once.

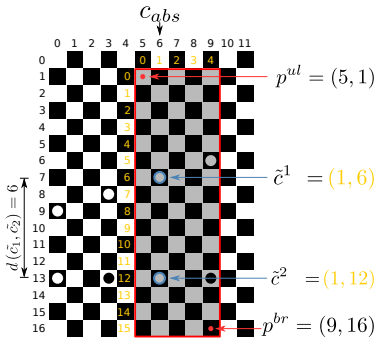


Fig. 9. Position estimation of the red marked unfolded subview with a dimension of $n_x = 4$ and $n_y = 15$. The absolute coordinates regarding the checkerboard pattern given in Fig. 4 are written in black. The relative coordinates are given in orange.

After the positions of the unfolded subviews are known, we combine the overlapping regions by using alpha-blending.

IV. Results

To demonstrate the quality and reliability of the proposed setup four aspects have been evaluated. First, we have examined the results of the cylindrical unfolding. This section is followed by an analysis of the performance of the system. In a third step we examined the image quality of in-vial FIM imaging. Finally we demonstrate the applicability of our setup for both, long term activity monitoring as well as posture and locomotion analysis.

A. Unfolding

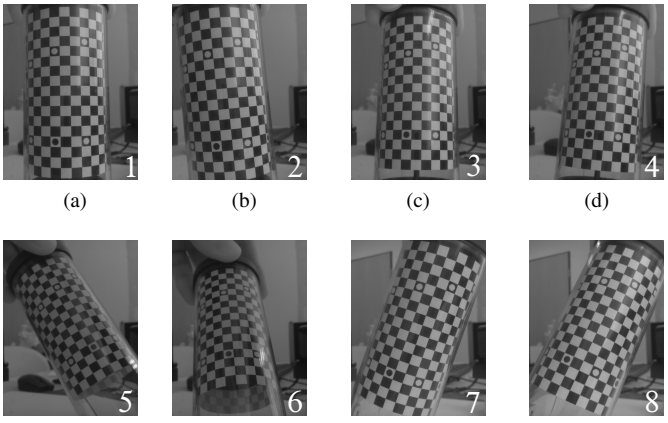
To evaluate the performance of the unfolding algorithm we have used 8 test images of a cylindrical vial with an inner radius of 22.63cm, a height of 11.2cm and a wall thickness of 2.4mm in different orientations with a checkerboard placed at the inner side of the vial. It should be noted that this system is not limited to this radius or height. The unfolding routine explained in Section III can be applied to arbitrary cylinder dimensions.

Since the same unfolding algorithm will run on each microcomputer simultaneously and since all 5 cameras are equal in terms of assemblage, we have used a randomly chosen one for this analysis. The testing images have a resolution of 992×1296 pixels and each of them is depicted in Fig. 10.

To evaluate the accuracy of the algorithm described in Section III-A and Section III-B we have used the re-projection error as an error metric. By knowing the parameters A, b, and f, the projected point q_i of a point $p_i \in \mathcal{M}$ on the image plane is given according to (4). Additionally, the associated detected keypoint $k_i \in \mathcal{K}$ as well as the interpolated point $\hat{q}_i$ are known. The re-projection error is thus given by $e = d\left(k_i, \hat{q}_i\right)$, where $d\left(k_i, \hat{q}_i\right)$ denotes the Euclidean distance between the image points k_i and $\hat{q}_i$. Furthermore, the mean re-projection error for all detected keypoints is given by

$$\left(\sum_{k_i \in \mathcal{K}} d\left(k_i, \hat{q}_i\right)\right) / |\mathcal{K}| \tag{5}$$

where $|\mathcal{K}|$ denotes the number of corners present in $\mathcal{K}$.



(a) (b) (c) (d)

(e) (f) (g) (h)

Fig. 10. (a-h) Images used for evaluation of the unfolding algorithm. The number of the images corresponds to the numbers used in Table I.

Each test was repeated 10 times. Table I summarizes the results where the mean error calculated by (5) of the 10 trials is shown. It can be seen that the mean error is always below 1 pixel. Compared to the approximate larva length of at least 15 pixels this accuracy is sufficient for our purpose.

TABLE I
MEAN NORMALIZED RE-PROJECTION ERROR OF THE TEST IMAGES.

Image	Mean Re-Projection Error (in px)
1	0.5285
2	0.5700
3	0.5713
4	0.5395
5	0.6532
6	0.8781
7	0.5432
8	0.6027

B. Performance

The stitching process can be performed quickly on the central server since all relevant parameters have to be calculated only once. Thus, this step is not the crucial part regarding the overall performance. To evaluate the performance of the system we measured the processing time for one transmission, i.e. we quantified the time of the following pipeline: First the server requests the unfolded images from all 5 clients. The clients then have to obtain new camera images in a certain resolution and subsequently perform the unfolding algorithm. Afterwards, the data is sent to the server. Once the server receives the data from all 5 clients, the transmission process is finished.

We have tested 6 different image resolutions, namely 544×960, 768×1024, 992×1296, 1088×1920, 1920×1904, and 2592×1936 in combination with 5 different square sizes $s \in \{20, 40, 60, 80, 100\}$, where s describes the size in pixels of a square edge in the unfolded image. The unfolded images had a size of $6s \times 16s$. Each transmission process was repeated 5000 times with a shutter speed of 100ms and the mean processing time as well as the standard deviation were

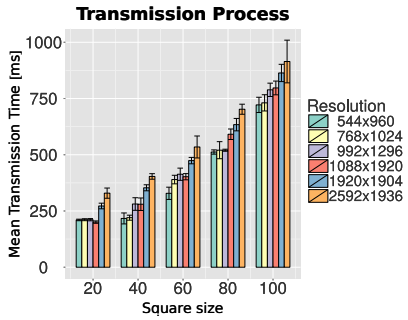


Fig. 11. Mean transmission processing time.

calculated. The results are summarized in Fig. 11 and clearly show that the transmission time strongly depends on both the image resolution and the square size. However, these factors have different influences. Looking at the values for $s = 20$ it can be seen that the resolution is not the dominant factor since the transmission time for the resolutions 544×960, 768×1024, 992×1296, and 1088×1920 is almost the same (approximately 220ms). Here the unfolding and the transmission through the network is the time-consuming part. For $s = 40$ and $s = 60$ the transmission time looks like expected. For a lower resolution and a lower s the process takes less time. Given a square size of $s = 80$ a similar observation can be made in comparison to a square size of $s = 20$. The processing time for the resolutions 544×960, 768×1024, and 992×1296 of approximately 500ms shows again that at this point the unfolding step is the elaborate part of the transmission process which prohibits a better performance of the system.

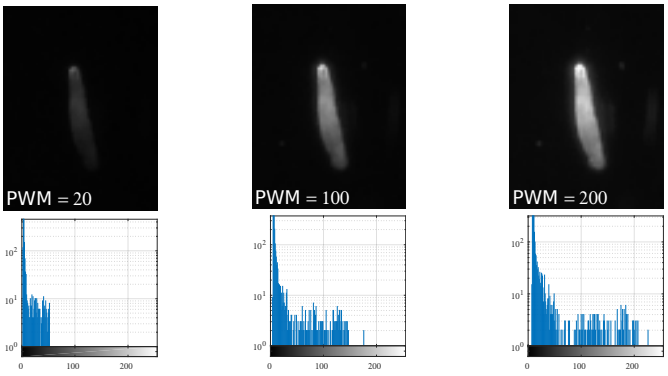
For our purpose a resolution of 992×1296 and a square size in the range of 30-40 represents an adequate trade-off between the accuracy and the performance of the system since a temporal resolution between 2 and 3 frames per second (fps) is sufficient for our purpose due to relatively slow movement of the larvae, even when computing higher level features such as body bending or area (see Section IV-E).

C. Contrast

A detailed comparison between images produced by the FIM system and images acquired by conventional imaging techniques has been given in [24]. Here we focused our analysis primarily on the contrast of the images produced by our setup. Fig. 12 provides an example of a Drosophila larva recorded with PWM values of 20, 100 and 200.

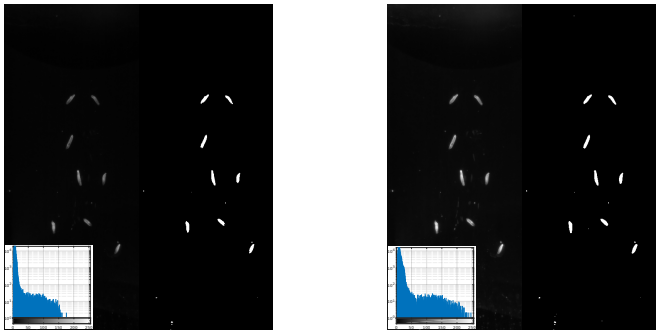
In grey-level images the contrast is the difference in luminance which makes objects distinguishable. The common definitions of contrast are based on ratios like $\frac{\text{luminance difference}}{\text{average luminance}}$ [35]. To measure the contrast we use grey-level histogram based variance analysis. In particular, Otsu's thresholding [36] is used since its terminology and parameters are well-established. Assuming bimodal grey-level distribution, the foreground and background can be assumed as two distinguishable Gaussian distributions with their maxima representing the mean foreground and background intensity. Since Otsu's method leads to the optimal threshold k in a bimodal distribution, this threshold can be used to separate the histograms as well as the associated

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBME.2016.2628203, IEEE Transactions on Biomedical Engineering



(a) (b) (c)

Fig. 12. Single larva captured with three different PWM values $(20, 100, 200)$. The histograms below are the logarithmic grey value distributions of the associated images.



(a) (b)

Fig. 13. Images recorded with PWM values of 100 (a) and 200 (b). The distributions of the background and the foreground are given in the associated histograms. Additionally, the binary images obtained by applying Otsu's thresholding are shown.

images in foreground and background regions. Furthermore, according to Otsu a measurement of separability is given by

$$\eta = \frac{\sigma_B^2}{\sigma_W^2 + \sigma_B^2}$$

where σ_W^2 and σ_B^2 are the within-class variance and the between-class variance, respectively. Fig. 13 shows two exemplary images recorded with PWM values of 100 and 200 as well as the histograms and the final segmentation using Otsu's thresholding.

The fact that the area of the inner vial surface is much higher than the total area covered by all visible larvae leads to an unequal foreground/background pixel distribution. Since a large, almost uniform, background is present, the average luminance is approximately equal to the background luminance. In such cases the *Weber contrast* is commonly used as a contrast quality measurement. The Weber contrast is part of the Weber-Fechner law and is given for grey-level images by

$$W_c = \frac{\mu_1 - \mu_0}{\mu_0}$$

where μ_0 and μ_1 are the mean luminances of the background and foreground.

To analyse the quality of the images and to explore the optimal PWM value range we acquired 5 images for each PWM value in the set of $\{10, 20, 30, 40, \ldots 240, 250\}$. The results for the 5 images of each PWM value were averaged to reduce the impact of image noise.

The curve depicted in Fig. 14a shows the relation between the PWM value and the optimal segmentation threshold k calculated by Otsu's algorithm. The slope of the curve is always positive and the changing rate of k is nearly a linear function of the PWM value which indicates that a good segmentation can be performed in a wide PWM range. This statement is supported by looking at Otsu's measurement of separability η. According to Otsu, η reaches its maximum at the minimal within-class variance. Thus this measurement can be used to find the optimal threshold and the optimal PWM value by plotting η against the PWM value. As illustrated in Fig. 14b the maximum value for η is reached at $PWM = 110$ with $\eta(110) = 0.763$. However, in the range of 70-200 the values for η differ from each other by maximum $\Delta\eta = 0.025$. Similarly, the optimal Weber contrast is reached at $PWM = 80$ with $W_c(80) = 14.955$ (Fig. 14c). Again the difference between all values of W_c in a range of 50-170 has a maximum of $\Delta W_c = 1.389$. If the PWM values for the maxima of η and W_c are exchanged the resulting measurements of $\eta(80) = 0.749$ and $W_c(110) = 14.464$ are still close to the respective maxima.

The results summarized in Fig. 14 suggest that a good segmentation of the foreground objects is possible. The best segmentation results are achieved with a PWM value range between 80 and 110 showing that our setup is not restricted to a specific parameter set but rather will produce adequate segmentation results without precise luminance adjustment.

Due to the fact that the used cylindrical vial is open on both sites (i.e. at the top and bottom), it is impossible to prevent a small amount of IR light from entering the inner part of the cylinder. This fact leads to a slightly inhomogeneous distribution of the background pixel values with growing PWM values and poses the question if high PWM values cause erroneous segmentation artefacts as such artefacts would increase the background class occurrence ω_0. By comparing all occurrences for the correctly segmented images (with the same number of larvae) and the given PWM values we quantify the maximum increase of noise in this distribution by

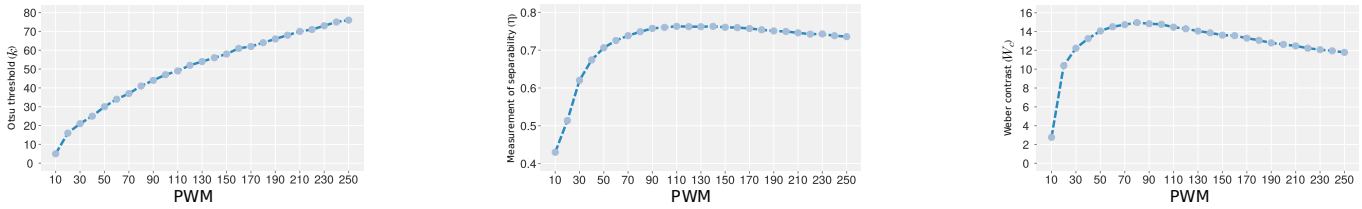
$$n_{inc} = \max_{p \in \{10,20,\ldots,250\}} \omega_0(p) - \min_{p \in \{10,20,\ldots,250\}} \omega_0(p)$$

Since $n_{inc} = 0.05809$, the range of pixels identified as background pixels is almost the same for all PWM values, our measurement technique is stable against IR light illumination changes.

D. Ghosts and Glowing Food

It was already mentioned in Section II that we were able to overcome the *ghost* and the *glowing food* problem by inserting an opaque inner cylinder in combination with an opaque food bowl. In Fig. 15 a three-dimensional representation of the cylindrical vial surface is given. It can be seen that there is almost no infrared light illuminating the inner part of the vial.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBME.2016.2628203, IEEE Transactions on Biomedical Engineering



(a) (b) (c)

Fig. 14. PWM study and contrast analysis. (a) Otsu's threshold k plotted against PWM. (b) Separability measurement η plotted against PWM. (c) Weber contrast W_c plotted against PWM.

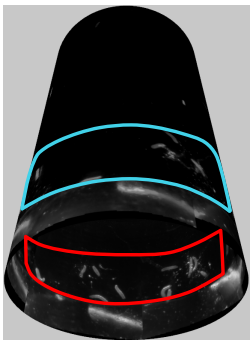


Fig. 15. Three-dimensional representation of the cylindrical vial surface. By inserting an opaque inner cylinder in combination with an opaque food bowl we were able to overcome the *ghost* and the *glowing food* problem.

Furthermore, the larvae marked red are not visible in the area marked blue and vice versa.

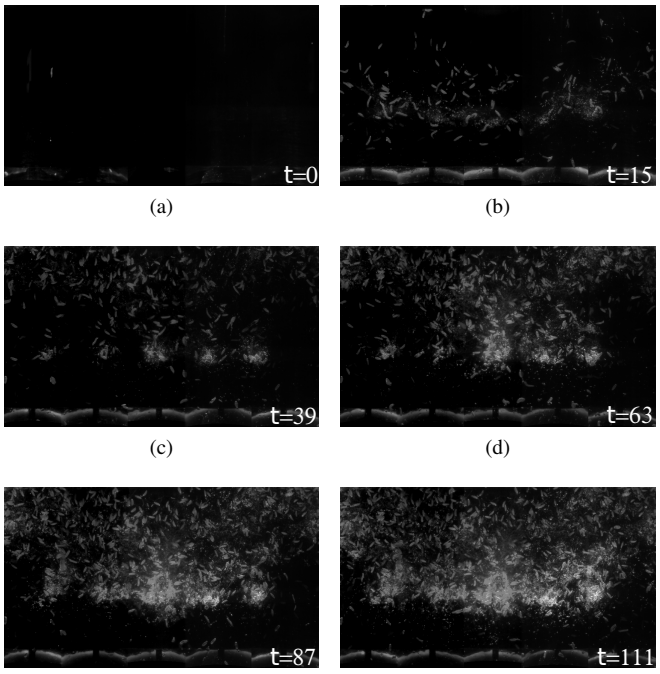
E. Locomotion Analysis

To evaluate the stability of our system and to demonstrate its abilities for long-term monitoring, we performed a long time test by capturing *Drosophila melanogaster* larvae 5 days after egg laying in a vial with an inner radius of 22.63cm, a height of 11.2cm, and a wall thickness of 2.4mm for 135 hours without interruption. During the recording the PWM value was set to 100 and according to the datasheet of the used LEDs this results in a radiant intensity of approximately $400\frac{mW}{sr}$.

To reduce the amount of data and to save storage space the capturing was done with 2 frames per second and both the stitched frames as well as the frames from each camera were stored in compressed *PNG* file format. Furthermore, images acquired by the cameras had a raw resolution of 992×1296 pixels. The square size s for unfolding was set to 30, resulting in an overall stitched image size of 840×480 pixels. In total, 560GB of data were acquired and saved for the subsequent evaluation.

Fig. 16 illustrates the first unfolded and stitched view of the vial where almost all larvae are located in the food at the food bowl. Additionally, images taken after $15, 39, 63, 87$ and 111 hours are depicted in the same Figure showing the distribution of the animals in the vial.

In the subsequent analysis the main challenge appears to be the elaboration of an adequate background model, since the dirt, the standard food, and the excretions of the larvae accumulate more and more on the tracking surface. Each of these artefacts induces the FTIR effect and as in any other optical imaging technique becomes apparent in the images. Additionally, these artefacts are not necessarily static since they can stick to larvae and can then be transferred to another position on the tracking surface very quickly (i.e. within a few frames). Furthermore, an increasing number of larvae start pupation over time, resulting in immoveable, static objects, which also induce the FTIR effect. Fig. 16 provides an impression of these problems and the development of the background over time. Modelling such a dynamic background where artefacts can have approximately the same size and even appearance as larvae is a challenging task.



(a) (b) (c) (d) (e) (f)

Fig. 16. (a-f) Unfolded and stitched raw images. The timestamps indicate the hour the images were taken at. The image at $t = 0$ is the first taken image.

To demonstrate the basic technical suitability of the proposed setup and to show the overall possibility to obtain meaningful readout, we excluded the background modelling problem by limiting the analysis to the first 15 hours. In total, we have analysed $107,671$ frames using FIMTrack [26] without any major adjustments for segmentation and tracking of the larvae. Currently, FIMTrack is not able to handle

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBME.2016.2628203, IEEE Transactions on Biomedical Engineering

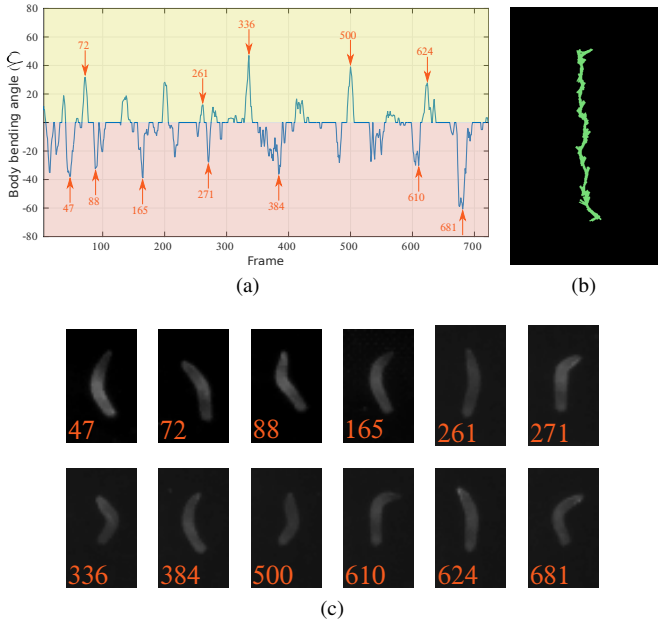


Fig. 17. (a) Body bending of a larva. The arrows correspond to the time points in the images in (c). (b) The complete track of the larva. (c) Images of the larva corresponding to the time points marked in (a).

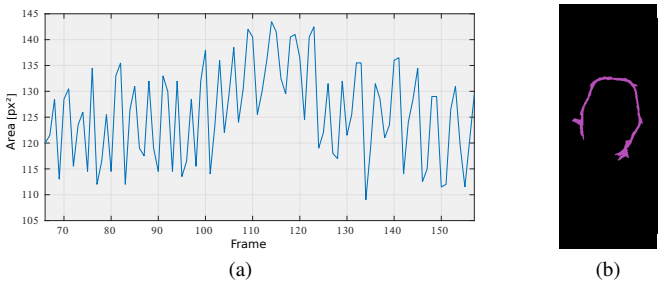


Fig. 18. (a) Area covered by the contour of a larva changes during the peristaltic contraction movement. (b) The associated trajectory of the larva.

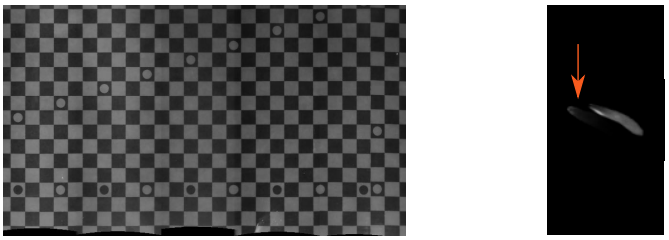


Fig. 19. Stitching artefacts: (a) unfolded and stitched checkerboard. (b) The artefact of a larva caused during the stitching process is marked by an arrow.

touching animals. Thus, these animals were excluded automatically from tracking and the associated trajectories were terminated at the time step at which the animals came in contact. Furthermore, we have discarded trajectories which are shorter than 30 seconds (i.e. 60 frames) resulting in $8,650$ remaining trajectories. Overall, we acquired more than 4.2×10^6 measurements. As mentioned above, this huge amount of data was acquired implicitly without both disturbing the animals and additional labour time.

As a first readout, we analysed the possibilities of FIMTrack to calculate posture and motion parameters like the body bending angle γ [26] or the area as an indirect peristalsis quantification parameter [24].

In Fig. 17a the body bending of a larva's trajectory given in Fig. 17b is plotted over time. A body bending to the left side is indicated by a positive value of γ (yellow area in Fig. 17a) and a bending to the right side is indicated by a negative value of γ (red area in Fig. 17a). The images of a single larva in Fig. 17c correspond to the time points marked by arrows in the body bending graph (Fig. 17a). In Fig. 17b it can be noted that utilizing FIMTrack for analysing images produced by the setup it is possible to obtain meaningful and analysable trajectories. Furthermore, FIMTrack is able to quantify the body bending of a larva. The quality of the images is still sufficient to detect even subtle body bending situations of a larva (see value and still image for time point 261 in Fig. 17a and Fig. 17c).

Moreover, the detected contour of a larva defines the area covered by the animal's body. During the peristaltic contraction movement, a change of the area is expected [37]. Fig. 18a illustrates the area of a larva's trajectory depicted in Fig. 18b. By analysing the changing rate of the area covered by the larval body it is possible to compute the contraction frequency of the animal.

V. Discussion

Imaging larvae in culture vials using several cameras and frustrated total internal reflection entails several challenges. We have already addressed and solved some of these. The *ghost* problem was resolved by inserting an inner opaque cylinder, effectively blocking scattered light from larvae at the opposite vial wall. Similarly, the *glowing food* was removed by disjoining the food and vial wall using a opaque food bowl.

One challenge occurs during the stitching process. In general, the calibration is done in relation to the inner surface of the cylinder described by the placement of the checkerboard. To facilitate a good stitching and blending, we assume that the position of the objects is the same in the overlapping area of both cameras, which is true given this surface (Fig. 19a). However, in some cases this assumption cannot be maintained. If a larva or a part of it moves too far away from the calibration surface, e.g. when the larva lifts its head or tail, the posture of this larva is imaged at two slightly different positions. Furthermore, this effect is aggravated by refraction effects. Since alpha blending is used, this may result in situations such as the one depicted in Fig. 19b where the artefact of a larva caused during the stitching process is marked by an arrow. It should be noted that the image of a larva is only influenced by this effect in a narrow area around the vertical stitching line. By modelling the refraction using a method similar to the one described in [38] and by applying a more sophisticated calibration and blending algorithm, this effect could potentially be decreased to a minimum.

The main challenge is caused by the accumulation of dirt during long measurements: Larvae leaving the food bowl slowly accumulate food and other artefacts at the vial surface. These artefacts also scatter IR light, making them detectable

in the images which aggravates a robust segmentation. Taking into account that the artefacts are mostly immoveable, this issue could be resolved by learning the background over time using a modified Gaussian mixture model approach [39]. Besides the accumulation of background artefacts the maximal monitoring time is also limited by the available disk space of the server. Given a well-defined behavioural readout online processing of the images would however strongly reduce the required memory load and could easily be incorporated due to the client server architecture of this system. Finally, the total amount of monitoring time has to be specified based on the specific experiment. Considering an analysis of wandering stage larvae a monitoring time of 24 hours might be sufficient. If differences in the initiation of the wandering stage or different hatching times need to be elucidated, longer monitoring periods are necessary.

Although using our setup it is possible to perform long-term monitoring of third instar Drosophila larvae without disturbing the animals while reducing the necessary effort of the experimenter to a minimum, currently it is not possible to perform stimulus-based behaviour analysis. However, a heat gradient could be integrated into the inner cylinder. Furthermore, since low-cost cameras are used the spatial resolution is not sufficient if very precise quantifications are necessary. This issue can be addressed by using more sophisticated but more expensive cameras which are used in systems like MAGAT [4] or FIM [24] in combination with more powerful hardware.

Nevertheless, in-vial tracking facilitates several novel studies. For example, this system represents a well-suited setup to analyse the molecular regulation of circadian rhythmicity. Several mutants of genes regulating the adult circadian rhythmic are known [40]. For the gene *timeless (tim)*, which is expressed from embryonic stages [40], different alleles are known to increase cycle frequency or to decrease it [41]. However, to our knowledge these studies so far have not been extended to the larval stage. Our system could allow to determine whether different locomotion strategies are used since it allows to track individual wandering third instar larvae after leaving the food for a long time period without disturbing them. Furthermore, this system for the first time allows an automated analysis of the negative geotropism shown by third instar larvae.

Additionally, the system may also provide a starting point for supervised Drosophila culture since it could potentially be extended by additional hardware to adjust specific environmental conditions like humidity in order to analyse if and how these adjustments influence the behaviour of the animals.

VI. Conclusion

We have proposed a novel method to monitor the behaviour of Drosophila larvae in vials from hours up to days. The system was used to image entire vials for up to 135 hours using multiple cameras and microcomputers. Our system provides both a suitable temporal resolution as well as a sufficient image resolution.

The subsequent posture and motion quantification was performed using FIMTrack [26], a software designed for analysing FIM images. We successfully analysed 15 hours of continuous recordings and computed both the trajectories and higher level features like body bending and the area of the larvae which can be used as an indirect measure for peristalsis. In total, $8,500$ tracks, each longer than half a minute, could be extracted, resulting in millions of data points. We are convinced that the proposed setup can be used for monitoring third intar crawling larvae, allowing to extract long-term quantities like time of entering the wandering stage, (day/night) activity, overall distribution in the vial, and time of pupation. Although we do not think that our system will replace all existing tracking paradigms, it has the potential to monitor the animals without any disturbance while drastically reduce labour time.

In the future we plan to improve both our background model and the calibration and stitching routine. Furthermore, additional algorithms like an automatic pupae classification and a collision resolution could be implemented to increase possible readouts. We are aiming for an online-tracking algorithm to extract the features during recording without saving the images which will reduce the required amount of disk space. This online-tracking in combination with a multi-vial imaging system in controlled environments will facilitate the extraction of billions of generic behavioural characteristics without changing the laboratory routine bringing a database of behavioural fingerprints of Drosophila larvae in reach.

Appendix A
Supplementary Movie

To give an impression of the images acquired by our setup we provide a supplementary movie with this document which shows a time laps of the first 39 hours of capturing *Drosophila melanogaster* larvae 5 days after egg laying.

Acknowledgment

The authors would like to thank J. Hermann for helping with the construction of the setup as well as A. Scherzinger for fruitful discussions throughout this project.

References

[1] N. Roussel *et al.*, "A Computational Model for *C. elegans* Locomotory Behavior: Application to Multiworm Tracking," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 10, pp. 1786–1797, Oct. 2007.

[2] K.-m. Huang *et al.*, "Automated Tracking of Multiple *C. Elegans* with Articulated Models," in *2007 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro*. IEEE, 2007, pp. 1240–1243.

[3] Y. Wang and B. Roysam, "Joint tracking and locomotion state recognition of *C. elegans* from time-lapse image sequences," *Biomedical Imaging: From Nano to Macro, 2010 IEEE International Symposium on*, pp. 540–543, 2010.

[4] M. Gershow *et al.*, "Controlling airborne cues to study small animal navigation," *Nature Methods*, vol. 9, no. 3, pp. 290–296, 2012.

[5] N. Otto *et al.*, "Interactions among Drosophila larvae before and during collision," *Scientific Reports*, vol. 6, p. 31564, 2016.

[6] Y. Xiang *et al.*, "Light-avoidance-mediating photoreceptors tile the *Drosophila* larval body wall," *Nature*, vol. 468, no. 7326, pp. 921–926, Dec. 2010.

[7] A. Gomez-Marin *et al.*, "Active sampling and decision making in *Drosophila* chemotaxis." *Nature Communications*, vol. 2, p. 441, 2011.

[8] T. A. Ofstad *et al.*, "Visual place learning in *Drosophila melanogaster*," *Nature*, vol. 474, no. 7350, pp. 204–207, 2011.

[9] K. Branson *et al.*, "High-throughput ethomics in large groups of *Drosophila*," *Nature Methods*, vol. 6, no. 6, pp. 451–457, Jun. 2009.

[10] N. A. Swierczek *et al.*, "High-throughput behavioral analysis in *C. elegans*," *Nature Methods*, vol. 8, no. 7, pp. 592–598, 2011.

[11] Z. Durisko *et al.*, "Dynamics of Social Behavior in Fruit Fly Larvae," *PLoS ONE*, vol. 9, no. 4, pp. 1–8, 2014.

[12] H. Jhuang *et al.*, "Automated home-cage behavioural phenotyping of mice," *Nature Communications*, vol. 1, p. 68, 2010.

[13] S. Ohayon *et al.*, "Automated multi-day tracking of marked mice for the analysis of social behaviour," *Journal of Neuroscience Methods*, vol. 219, no. 1, pp. 10–19, 2013.

[14] A. Weissbrod *et al.*, "Automated long-term tracking and social behavioural phenotyping of animal colonies within a semi-natural environment," *Nature Communications*, vol. 4, p. 2018, 2013.

[15] S. H. Simonetta and D. A. Golombek, "An automated tracking system for *Caenorhabditis elegans* locomotor behavior and circadian studies application," *Journal of Neuroscience Methods*, vol. 161, no. 2, pp. 273–280, 2007.

[16] C.-C. J. Yu *et al.*, "Multi-well imaging of development and behavior in *Caenorhabditis elegans*," *Journal of Neuroscience Methods*, vol. 223, pp. 35–39, 2014.

[17] S. Nagy *et al.*, "Measurements of behavioral quiescence in *Caenorhabditis elegans*," *Methods*, vol. 68, no. 3, pp. 500–507, 2014.

[18] H. Dankert *et al.*, "Automated monitoring and analysis of social behavior in *Drosophila*," *Nature Methods*, vol. 6, no. 4, pp. 297–303, 2009.

[19] T. Yoshii *et al.*, "Cryptochrome Mediates Light-Dependent Magnetosensitivity of *Drosophila*'s Circadian Clock," *PLoS Biol*, vol. 7, no. 4, pp. 1–7, 04 2009.

[20] O. T. Inan *et al.*, "A Miniaturized Video System for Monitoring the Locomotor Activity of Walking *Drosophila Melanogaster* in Space and Terrestrial Settings," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 2, pp. 522–524, Feb 2009.

[21] G. F. Gilestro, "Video tracking and analysis of sleep in *Drosophila melanogaster*," *Nature Protocols*, vol. 7, no. 5, pp. 995–1007, 2012.

[22] D. S. Garbe *et al.*, "Context-specific comparison of sleep acquisition systems in *Drosophila*," *Biology open*, vol. 4, no. 11, pp. 1558–1568, 2015.

[23] TriKinetics, "TriKinetics: locomotion Activity Monitoring system for biological research." http://www.trikinetics.com/, 2016.

[24] B. Risse *et al.*, "FIM, a Novel FTIR–Based Imaging Method for High Throughput Locomotion Analysis," *PLoS ONE*, vol. 8, no. 1, p. e53963, Jan. 2013.

[25] B. Risse *et al.*, "Quantifying subtle locomotion phenotypes of *Drosophila* larvae using internal structures based on FIM images," *Computers in Biology and Medicine*, vol. 63, pp. 269–276, 2015.

[26] B. Risse *et al.*, "FIM Imaging and FIMTrack: Two New Tools Allowing High-throughput and Cost Effective Locomotion Analysis," *Journal of Visualized Experiments: JoVE*, no. 94, p. e52207, 2014.

[27] B. Risse *et al.*, "FIM^{2c} : A Multi-Colour, Multi-Purpose Imaging System to Manipulate and Analyse Animal Behaviour," *IEEE Transactions on Biomedical Engineering*, 2016, in press.

[28] G. Slater *et al.*, "A Central Neural Pathway Controlling Odor Tracking in *Drosophila*," *The Journal of Neuroscience*, vol. 35, no. 5, pp. 1831–1848, 2015.

[29] M. Q. Clark *et al.*, "Functional Genetic Screen to Identify Interneurons Governing Behaviorally Distinct Aspects of Drosophila Larval Motor Programs," *G3: Genes—Genomes—Genetics*, vol. 6, no. 7, pp. 2023–2031, 2016.

[30] H. Bach and N. Neuroth, *The properties of optical glass*. Springer Science & Business Media, 2012.

[31] Raspberry Pi, "Source code for ARM side libraries for interfacing to Raspberry Pi GPU," https://github.com/raspberrypi/userland, 2016.

[32] W. Puech *et al.*, "Projection distortion analysis for flattened image mosaicing from straight uniform generalized cylinders," *Pattern Recognition*, vol. 34, no. 8, pp. 1657–1670, 2001.

[33] Z. Ye *et al.*, "Reading labels of cylinder objects for blind persons," in *2013 IEEE International Conference on Multimedia and Expo (ICME)*, 2013, pp. 1–6.

[34] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed. New York: Cambridge University Press, 2003.

[35] E. Peli, "Contrast in complex images," *Journal of the Optical Society of America A*, vol. 7, no. 10, pp. 2032–2040, 1990.

[36] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.

[37] S. Lahiri *et al.*, "Two Alternating Motor Programs Drive Navigation in *Drosophila* Larva," *PLoS ONE*, vol. 6, no. 8, pp. 1–12, 08 2011.

[38] S. Schmid, "Development and Evaluation of a Tracking System for Imaging Freely-Moving Mice inside the quadHIDAC PET Scanner," Ph.D. dissertation, University of Münster, 2016.

[39] N. Greggio *et al.*, "Self-Adaptive Gaussian Mixture Models for Real-Time Video Segmentation and Background Subtraction," in *2010 10th International Conference on Intelligent Systems Design and Applications*, 2010, pp. 983–989.

[40] B. R. Crane and M. W. Young, "Interactive Features of Proteins Composing Eukaryotic Circadian Clocks," *Annual review of biochemistry*, vol. 83, pp. 191–219, 2014.

[41] A. Rothenfluh *et al.*, "Isolation and Analysis of Six *timeless* Alleles That Cause Short- or Long-Period Circadian Rhythms in Drosophila," *Genetics*, vol. 156, no. 2, pp. 665–675, 2000.